

Recovering the Brownian Coalescent Point Process from the Kingman Coalescent by Conditional Sampling

Amaury Lambert, Emmanuel Schertzer

November 7, 2016

Abstract

We consider a continuous population whose dynamics is described by the standard stationary Fleming-Viot process, so that the genealogy of n uniformly sampled individuals is distributed as the Kingman n -coalescent. In this note, we study some genealogical properties of this population when the sample is conditioned to fall entirely into a subpopulation with most recent common ancestor (MRCA) shorter than ε . First, using the comb representation of the total genealogy [LUB16], we show that the genealogy of the descendance of the MRCA of the sample on the timescale ε converges as $\varepsilon \rightarrow 0$. The limit is the so-called Brownian coalescent point process (CPP) stopped at an independent Gamma random variable with parameter n , which can be seen as the genealogy at a large time of the total population of a rescaled critical birth-death process, biased by the n -th power of its size. Secondly, we show that in this limit the coalescence times of the n sampled individuals are i.i.d. uniform random variables in $(0, 1)$. These results provide a coupling between two standard models for the genealogy of a random exchangeable population: the Kingman coalescent and the Brownian CPP.

1 Introduction

In this paper, we seek to couple two well-known probabilistic objects both modeling the genealogy of an exchangeable population. The first object is the celebrated *Kingman n -coalescent* [K82], which arises as the genealogy of a sample of n co-existing individuals within a stationary population of large size N constant through time, when time is measured in units of N generations. The second object is the *coalescent point process* introduced in [P04], which arises as the genealogy of the whole population at time N of a critical birth-death process starting from size N and with birth and death rates both equal to 1, when time is also accelerated by N .

In the coalescent point process, the population is assumed to be endowed with a linear order consistent with the genealogy, in the sense that in a plane representation of this genealogy, lineages only intersect at internal nodes (common ancestors) –see Fig. 1 or Fig. 7 in [L08]. This order can also be obtained as the order inherited from a contour of the tree [P04, L10, LP13]. The linear arrangement of coalescence times (that is, times to the MRCA –most recent common ancestor) between consecutive pairs of individuals ranked in the linear order converges to the concatenation of (a Poisson number with parameter 1 of) i.i.d. Poisson point processes with intensity measure $2dt x^{-2}dx$ killed at their first atom with second coordinate larger than 1. Each of these killed Poisson point processes encodes the genealogy

of the descendance of an individual in a critical branching process conditioned on survival up to a large time. They will hereafter be called killed Brownian coalescent point process (killed Brownian CPP).

Both the standard Kingman coalescent and the killed Brownian CPP code for the genealogy of a large exchangeable population, but there are two features distinguishing them. First, the Kingman coalescent focuses on sparsely sampled individuals whereas the CPP deals with the whole population. Second, the Kingman coalescent is based on the assumption of a stationary population with constant size (total size constraint), whereas in the CPP the size of the population is only constant in expectation, and its foundation time is fixed (time constraint).

Rather than defining a third object coupling the Kingman coalescent and the CPP, our aim is to show that one of the two is embedded in the other. Due to the first aforementioned distinctive feature (i.e., sparse sampling), one might think at first sight that the Kingman coalescent can be obtained by sparsely sampling the CPP. But in doing this, one would not get rid of the second distinctive feature, namely the time constraint. The alternative view is the right one. In an exchangeable population with large constant size, the descendance of a small subpopulation is blind to the total size constraint and it is constant in expectation (see for example Theorem 1 in [BL06]). Our goal is to prove a backward-in-time version of the last informal statement (‘in a large stationary population with constant size, the genealogy of a subpopulation with recent MRCA is given by a CPP’) and to derive some consequences of this fact.

We start from the genealogy of a population with constant size in the stationary case directly with the continuous limiting object, the standard Fleming-Viot process [FV79]. Actually, we will make use of an alternative description of the Fleming-Viot process, namely the flow of bridges introduced by Bertoin and Le Gall [BL03], in which the population is endowed with a linear order consistent with the genealogy (See Section 2 and Fig. 1). We will call *Kingman comb* the list of coalescence times of pairs of ‘consecutive individuals’ in this linearly arranged continuous population, as defined in Section 3. In Section 3, we show that the properly rescaled Kingman comb converges to the Brownian CPP (see Proposition 3.4).

In the remainder of the paper, we investigate the genealogy of n individuals sampled in such a way that their MRCA lies at a depth smaller than ε with $\varepsilon \rightarrow 0$. Note that this conditioning can be implemented in two distinct ways:

- (i) (*Quenched conditional sampling*) Conditional on the flow of bridges, sample n individuals such that their MRCA lies at a depth smaller than ε and then average over every realization of the flow. In biology, such conditioning could arise by sampling on purpose individuals that share close phenotypic characteristics or dwell in neighboring habitats.
- (ii) (*Averaged conditional sampling*) Directly condition the n -Kingman coalescent to have its MRCA lie at a depth smaller than ε . In contrast with (i), where the conditioning is solely enforced at the sampling level, the conditioning in (ii) could be due either to anomalous sampling (as in (i)) or to an abnormally shallow MRCA of the *entire* population.

In Section 4, we focus on case (i), where we consider the entire family that shares a common ancestor with the n sampled individuals. We show that for small ε , the genealogy of this family is given by a CPP killed at an independent Gamma random variable (see Theorem 4.1). Informally, this amounts to saying that the genealogy of the family of the sample is

the rescaled genealogy of a *n-size-biased* critical birth-death process (i.e., biased by the n -th power of its size) conditioned on survival up to a large time (see Remark 4.2).

In Section 5.1, we use this result to prove that the genealogy of the n conditionally sampled individuals enjoy a nice i.i.d. structure, namely that their (properly rescaled) $n-1$ coalescence times are i.i.d. uniform random variables (Theorem 5.1). In Section 5.2, we turn to case (ii) where we show (Theorem 5.4) that the genealogy of the n individuals is also described (asymptotically) in terms of i.i.d. uniform random variables, thus showing that the conditionings (i) and (ii) become indistinguishable as ε goes 0.

Finally, we briefly mention a natural conjecture arising from the previous results. Because (a) the genealogy of the n individuals in (i) and (ii) coincide asymptotically as $\varepsilon \rightarrow 0$, and because (b) the entire family sharing a common ancestor with the n individuals in case (i) is described in terms of a ‘ n -size-biased’ killed Brownian CPP (see again Theorem 4.1), it is natural to conjecture that Theorem 4.1 also holds in case (ii).

2 Preliminaries: Flows of Bridges and Combs

2.1 Discrete Bridges

Flows of bridges have been introduced by Bertoin & Le Gall in [BL03]. In order to motivate their construction, let us consider a general discrete time Cannings [C75] model as follows.

- (1) At each generation the size of the population is fixed and equal to N ;
- (2) Individuals at generation r are labelled from 1 to N and we denote by $(\nu_r^1, \dots, \nu_r^N)$ the vector of offspring numbers;
- (3) This labeling of individuals is consistent with the genealogy (cf. Introduction and Fig. 1);
- (4) The vectors $((\nu_r^1, \dots, \nu_r^N); r \in \mathbb{Z})$ are i.i.d. exchangeable vectors.

Recall from the Introduction that a labeling consistent with the genealogy is such that lines of descent only cross at internal nodes (See Fig. 1). Rigorously, this amounts to enforcing the condition that if $i < j$, the label of an offspring of individual i is always smaller than the label of an offspring of individual j .

Now for any $x \in \{0, \dots, N\}$ and $r \in \mathbb{Z}$, define $B_{r,r+1}(x)$ as the number of individuals at generation $r+1$ descending from the subpopulation with labels smaller than or equal to x at generation r . Thanks to Assumption (2), we have

$$B_{r,r+1}(x) := \sum_{k \leq x} \nu_r^k.$$

Thanks to Assumptions (1) and (4), the maps $(B_{r,r+1}; r \in \mathbb{Z})$ are i.i.d. and each $B_{r,r+1}$ is a *discrete bridge*, that is a non-decreasing function from $\{0, \dots, N\}$ onto itself with exchangeable increments. For any $m < n$, define

$$B_{m,n} := B_{n-1,n} \circ \dots \circ B_{m,m+1}.$$

Thanks to Assumption (3), $B_{m,n}(x)$ is the number of individuals at generation n descending from the subpopulation with labels smaller than x at generation m . The bridge property is stable under composition and furthermore, $\{B_{m,n}\}_{m < n}$ satisfies the so-called cocycle property

$$B_{k,n} = B_{m,n} \circ B_{k,m} \quad k < m < n.$$

The collection $\{B_{m,n}\}_{m < n}$ is called a discrete *flow* of discrete bridges. Clearly, the increments of the flow are stationary and independent.

Let us define the inverse flow $\phi_{m,n}$ by

$$\phi_{m,n}(x) = \inf\{y \in \{1, \dots, N\} : B_{m,n}(y) \geq x\} \quad m < n, \quad x \in \{1, \dots, N\}. \quad (1)$$

The bridge property implies that $\{\phi_{m,n}\}_{m < n}$ defines a backward coalescing flow, in the sense that $\phi_{k,n} = \phi_{k,m} \circ \phi_{m,n}$ for $k < m < n$ and the orbits $\{\phi_{m,n}(x)\}_{m \leq n}$ coalesce upon meeting each other as m decreases. More specifically, if $\phi_{m,n}(x) = \phi_{m,k}(x')$, then $\phi_{m',n}(x) = \phi_{m',k}(x')$ for all $m' < m$. For every $x \in \{1, \dots, N\}$, we can then see the orbit

$$\begin{aligned} \mathbb{N} &\rightarrow \{1, \dots, N\} \\ k &\mapsto \phi_{n-k,n}(x) \end{aligned}$$

as the *ancestral lineage* of individual x of generation n . (See Fig. 1.)

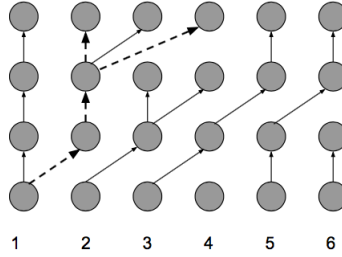


Figure 1: A realization of the Cannings model at times $t = 0, 1, 2, 3$ flowing upwards. Dashed lines indicate the ancestral lineages of individuals 2 and 4 at generation 3. At each generation, individuals are labelled from 1 to 6 so that ancestral lines do not cross.

2.2 The Standard Fleming-Viot Flow of Bridges

In a similar way to the discrete flow of bridges, we now define the continuous flow of bridges as done in [BL03]. Now a *bridge* is a non-decreasing function B from $[0, 1]$ onto itself with exchangeable increments such that $B(0) = 0$ and $B(1) = 1$. A stochastic *flow of bridges* is a family $\{B_{s,t}\}_{-\infty < s < t < \infty}$ of bridges satisfying the following properties.

- (1) Co-cycle property. For any fixed $r < s < t$, $B_{r,t} = B_{s,t} \circ B_{r,s}$ a.s.;
- (2) Independent and stationary increments;
- (3) No fixed time discontinuity. For any fixed time s , $\lim_{t \downarrow s} B_{s,t} = \text{Id}$ (uniformly) in probability.

We think of a stochastic flow as the dynamics of a stationary, continuous population with constant size equal to 1. The genealogy of the population alive at time t is encoded by the backward coalescing flow $\{\phi_{s,t}\}_{s < t}$ defined analogously to (1) by

$$\phi_{s,t}(x) = \inf\{y \in [0, 1] : B_{s,t}(y) > x\} \quad s < t, \quad x \in [0, 1]. \quad (2)$$

Since the increments of a flow are stationary and independent, a flow is uniquely characterized by its one-dimensional marginal $B_{0,t}$. In this paper, we will specifically consider the so-called *standard Fleming-Viot* (FV) flow of bridges whose one-dimensional marginal is equal to

$$B_{0,t}(x) = \sum_{i=1}^{N_t} \mathbb{1}_{[0,x]}(U_i) \beta_i$$

where

- (1) N_t is distributed as the value at time t of a pure-death process going from k to $k-1$ at rate $k(k-1)/2$ and started at ∞ ;
- (2) Conditional on $N_t = n$, the random vector $\{\beta_i\}_{i=1}^n$ is independent of the (U_i) and follows the Dirichlet distribution with parameter $(1, \dots, 1)$;
- (3) $\{U_i\}_{i=1}^\infty$ is a sequence of i.i.d. uniform random variables independent of $(N_t; \{\beta_i\}_{i=1}^{N_t})$.

In the same spirit as [MS01], this flow should arise as the scaling limit of discrete bridges induced by any Cannings model with enough control on the tail of the offspring distribution (in particular the Wright-Fisher model).

2.3 Combs and Coalescent Point Processes

In the next section, we will be interested in the backward coalescing flow associated with the FV flow of bridges. The precise trajectory of the ancestral lineage of a given individual – i.e., the successive labels in $(0,1)$ of the ancestors of this individual – is in most applications irrelevant. In contrast, one would like to extract from the coalescing flow the pure genealogical information contained in this object.

To do this we follow [LUB16] and define a *comb* as a function $f : [0, +\infty) \rightarrow [0, +\infty)$, such that for any $\varepsilon > 0$, $\{f \geq \varepsilon\}$ is finite. Then $d_f(x, y) := \sup_{z \in (x \wedge y, x \vee y)} f(z)$ defines an ultrametric distance on $\{f = 0\}$ called the comb metric (modulo the identification of points at distance 0 if f is zero on one or several open intervals). Let Ω be the space of combs and consider the mapping

$$\begin{aligned} F : \Omega &\rightarrow \mathcal{M} \\ f &\mapsto \sum_{x: f(x) > 0} \delta_{x, f(x)} \end{aligned}$$

where \mathcal{M} denotes the space of point measures on $(0, \infty)^2$. We assume that \mathcal{M} is endowed with the topology induced by test-functions $\varphi : (0, \infty)^2 \rightarrow \mathbb{R}$ which are continuous and bounded and for which there is $M > 0$ and $\varepsilon > 0$ such that $\varphi = 0$ outside $(0, M) \times (\varepsilon, \infty)$. We equip Ω with the σ -field generated by F when \mathcal{M} is equipped with its Borel σ -field.

Let (C_n) be a sequence of random combs such that for every $A \subseteq (0, \infty)^2$ with zero Lebesgue measure, $F(C_n)(A) = 0$ a.s. for every n . Let C be a random comb with the same property. We will use repeatedly the fact that (C_n) converges weakly in law to C iff for any $x_1 < \dots < x_k$, the random vectors $(d_{C_n}(x_i, x_{i+1}); 1 \leq i \leq k-1)$ converge in law to the random vector $(d_C(x_i, x_{i+1}); 1 \leq i \leq k-1)$. This can be seen thanks to the Portmanteau theorem and thanks to the following equality between events

$$\{d_C(x_i, x_{i+1}) < y_i, \forall i\} = \{F(C)(A) = 0\},$$

where A is the following subset of $(0, \infty)^2$

$$A = \cup_{i=1}^{k-1} (x_i, x_{i+1}) \times [y_i, \infty),$$

which is such that $F(C)(\partial A) = 0$ a.s.

For any comb f and $t > 0$, we will set

$$k_t(f) := f \mathbb{1}_{[0,t)}$$

and call it the comb f *killed at t* . For any $x > 0$, we define

$$l_x(f) := \inf\{y \geq 0 : f(y) > x\}.$$

For any $a > 0$, we define the scaling operator S_a by

$$S_a(f)(x) := a^{-1} f(a^{-1}x).$$

In particular, $S_a \circ k_t = k_{t/a} \circ S_a$ and $S_a \circ k_{l_x} = k_{l_{x/a}} \circ S_a$.

For any σ -finite measure ν on $(0, \infty)$, the *coalescent point process* (CPP) with intensity measure ν is the random comb C such that $F(C)$ is a Poisson point process with intensity measure $dt \otimes \nu(dx)$. The coalescent point process associated to

$$\nu(dx) = \frac{2}{x^2} dx,$$

will hereafter be denoted \mathcal{C} and called the *Brownian CPP*. We will also define $\bar{\mathcal{C}} := k_{l_1}(\mathcal{C})$ referred to as the *killed Brownian CPP*. As mentioned in the introduction, $\bar{\mathcal{C}}$ can be thought of as the genealogy rescaled by N of the descentance of an individual by a critical branching process conditioned on survival up to a large time N .

Remark 2.1. *Based for example on [LG05], it is known that the reflected Brownian motion codes in a certain appropriate sense for a rescaled critical branching forest. For a forest coded by a non-negative function h , the coalescence times of the part of the tree lying at distance d from the root, are the depths of the excursions of h away from d . This explains why the measure ν is (up to a multiplicative constant) the Itô measure of Brownian excursion depths [P04].*

3 The Kingman Comb at Small Scale

3.1 The Kingman Comb

Let ϕ be the backward coalescing flow defined in (2) from the standard Fleming-Viot flow of bridges. For every $x < y \in (0, 1)$, the coalescence time of x and y is given by

$$\inf\{t > 0 : \phi_{-t,0}(x) = \phi_{-t,0}(y)\}.$$

The next statement shows that this genealogical structure can be represented as a comb that we call the Kingman comb (see also [K82, LUB16] for other treatments of the Kingman comb).

Proposition 3.1. *There is a sequence (V_j) of i.i.d. uniform random variables and an independent sequence (T_j) where $T_j = \sum_{k \geq j+1} e_k$, for e_k independent exponential r.v. with parameter $k(k-1)/2$, such that*

$$\inf\{t > 0 : \phi_{-t,0}(x) = \phi_{-t,0}(y)\} = d_C(x, y) \quad \text{for every } x < y \text{ in } \{C = 0\} \text{ a.s.}$$

where

$$C = \sum_{j \geq 1} T_j \mathbb{1}_{\{V_j\}}. \quad (3)$$

The function C is a comb a.s., and the distance d_C is the comb metric associated to C . Thereafter, we will call this random comb the Kingman comb.

Proof. The fact that C is a comb is straightforward. Instead, we focus on the first part of Proposition 3.1. Let (W_n) be an independent sequence of i.i.d. uniform r.v. and for any $t > 0$ define the equivalence relation \sim_t in \mathbb{N} by

$$k \sim_t n \iff \phi_{-t,0}(W_k) = \phi_{-t,0}(W_n).$$

We denote by Π_t the partition of \mathbb{N} induced by \sim_t . It is known from [BL03] that $(\Pi_t; t \geq 0)$ (has a càdlàg modification which) is distributed as the standard Kingman coalescent. In particular, the number of blocks N_t of Π_t is a non-increasing process started at ∞ which jumps from k to $k-1$ at rate $k(k-1)/2$, so that the intersection R_t of $(0, 1)$ with the range of $B_{-t,0}$ is finite (with cardinal $N_t - 1$) and non-increasing. Let $J_1 > J_2 > \dots$ denote the jump times of (N_t) labelled in decreasing order and for any $i \geq 1$ let V_i be the unique element of $(0, 1)$ such that

$$R_{J_i} = R_{J_{i-1}} \cup \{V_i\}$$

We also know that $B_{-t,0}$ can be written as

$$B_{-t,0}(x) = \sum_{i=1}^{N_t} \mathbb{1}_{[0,x]}(U_i) \beta_i$$

where conditional on $N_t = n$, the vector (β_i) follows the Dirichlet distribution with parameter $(1, \dots, 1)$. This means that for all n , the vector of diameters of the connected components of $(0, 1) \setminus \{V_1, \dots, V_{n-1}\}$ follows the Dirichlet distribution with parameter $(1, \dots, 1)$. Standard arguments imply that the (V_i) form a sequence of i.i.d. uniform r.v. independent of the sequence (W_n) (because they depend deterministically upon the flow of bridges).

Next define $e_k := J_{k-1} - J_k$ and $T_j = \sum_{k \geq j+1} e_k$. We have already mentioned that e_k is exponentially distributed with parameter $k(k-1)/2$ and it is well-known that the (e_k) are independent of the asymptotic frequencies of the blocks of (Π_t) , and so are independent of the (V_i) . Defining C as in (3), we easily see that a.s. for any k, n , for any $t \in [J_i, J_{i-1})$ ($i \geq 1$, $J_0 := +\infty$),

$$k \sim_t n \iff \forall j \in \{1, \dots, i-1\}, V_j \notin (W_k \wedge W_n, W_k \vee W_n) \iff \sup_{(W_k \wedge W_n, W_k \vee W_n)} C < t.$$

Taking the union of i , this can be expressed as follows. Almost surely for any k, n , for any t which is not a jump time of N ,

$$\phi_{-t,0}(W_k) = \phi_{-t,0}(W_n) \iff \sup_{(W_k \wedge W_n, W_k \vee W_n)} C < t.$$

By density of the (W_n) , this implies that a.s. for all $x < y \in (0, 1) \setminus \{U_i\}$, for all t which is not a jump time of N ,

$$\phi_{-t,0}(x) = \phi_{-t,0}(y) \iff \sup_{(x,y)} C < t,$$

which terminates the proof. ■

Remark 3.2. *Proposition 3.1 is interesting in its own right. It provides a natural interpretation of the r.v. V_i appearing in the definition of the Kingman comb. Thinking of the flow of bridges as describing the dynamics of a population of constant size 1, the T_i 's indicate the dates of branching events giving rise to lineages surviving both until the present. For a given value of T_i , the two resulting extant subpopulations can be identified with the interval (V_i^-, V_i) and (V_i, V_i^+) where*

$$\begin{aligned} V_i^+ &= \inf\{V_j > V_i : T_j > T_i\}, \\ V_i^- &= \sup\{V_j < V_i : T_j > T_i\}, \end{aligned}$$

with the convention $\inf\{\emptyset\} = 0$ and $\sup\{\emptyset\} = 1$. To conclude, not only does the comb encapsulate the time and the linear ordering of splitting events that are relevant to the present (the T_i 's), but it also retains the size of the sub-populations arising from those splitting events (the V_i 's).

Remark 3.3. *The comb at time $t = 0$ is generated from the sequence $(\phi_{-t,0}; t \geq 0)$. Analogously, for every time $s \in \mathbb{R}$, one can define a comb C_s from the sequence $(\phi_{t,s}; t \leq s)$. $(C_s; s \in \mathbb{R})$ naturally defines a stationary stochastic process that will be the subject of future work. We expect that this ‘evolving Kingman comb’ will shed new light on the evolving Kingman coalescent studied for example in [PW06] and [PWW11].*

3.2 Convergence to the Brownian CPP

The next proposition relates the Kingman comb at small scale with the Brownian CPP.

Proposition 3.4. *The following convergence $S_\varepsilon(C) \implies \mathcal{C}$ holds weakly in law as $\varepsilon \rightarrow 0$.*

Proof. Let $\varepsilon > 0$. For every $t_0 > 0$ and every bounded open interval $I \subseteq (0, \infty)$, define

$$F_\varepsilon^{I,t_0} = \{(u_\varepsilon, s_\varepsilon) \in (t_0, \infty) \times I : S_\varepsilon(C)(u_\varepsilon) = s_\varepsilon\}.$$

In order to prove Proposition 3.4, it is enough to show that the random set F_ε^{I,t_0} converges weakly in the vague topology to a PPP on $(0, \infty)^2$ with intensity measure

$$2 \, 1_{x \in I} dx \, 1_{t \geq t_0} \frac{dt}{t^2}$$

as $\varepsilon \rightarrow 0$. For every $t > 0$, define $N_t := \#\{j : T_j > t\}$ corresponding to the block counting process of the Kingman coalescent. By definition, the set F_ε^{I,t_0} coincides with the set of points of the form $\frac{1}{\varepsilon}(V_i, T_i)$ with (V_i, T_i) belonging to the Kingman comb (as defined in Proposition 3.1) and such that

$$V_i \in \varepsilon I \text{ and } i \in \{1, \dots, N_{\varepsilon t_0}\}.$$

Next, let $c > 0$ and let us compare the previous set with the set $\bar{F}_\varepsilon^{I,c}$ consisting of every point of the form $\frac{1}{\varepsilon}(V_i, T_i)$, with (V_i, T_i) again belonging to the Kingman comb, but such that

$$V_i \in \varepsilon I \quad \text{and} \quad i \in \{1, \dots, [\frac{c}{\varepsilon}]\}.$$

We claim that $\bar{F}_\varepsilon^{I,c}$ converges to a PPP with intensity measure

$$2 \, 1_{x \in I} dx \, 1_{t \geq \frac{2}{c}} \frac{dt}{t^2}. \quad (4)$$

Before justifying the claim, let us briefly explain how this entails Proposition 3.4. On the one hand, for any $c_1 < c_2$

$$\mathbb{P} \left(\bar{F}_\varepsilon^{I,c_1} \leq F_\varepsilon^{I,t_0} \leq \bar{F}_\varepsilon^{I,c_2} \right) \geq \mathbb{P} \left(N_{\varepsilon t_0} \in \left\{ [\frac{c_1}{\varepsilon}], \dots, [\frac{c_2}{\varepsilon}] \right\} \right).$$

It is well known that the renormalized block counting process $\varepsilon t_0 N_{\varepsilon t_0}$ converges to 2 in probability as $\varepsilon \downarrow 0$. Thus, for any $\delta > 0$, taking $c_1 = \frac{2-\delta}{t_0}$ and $c_2 = \frac{2+\delta}{t_0}$ in the latter inequality yields

$$\mathbb{P} \left(\bar{F}_\varepsilon^{I,c_1} \leq F_\varepsilon^{I,t_0} \leq \bar{F}_\varepsilon^{I,c_2} \right) \rightarrow 1 \quad \text{as } \varepsilon \rightarrow 0.$$

Finally, since this holds for every $\delta > 0$, Proposition 3.4 easily follows by letting $\delta \rightarrow 0$ (assuming that $\bar{F}_\varepsilon^{I,c}$ converges to a PPP with the intensity measure provided in (4)).

It remains to justify the convergence of $(\bar{F}_\varepsilon^{I,c})$. Recall from Proposition 3.1 that the r.v. V_1, V_2, \dots in the Kingman comb are i.i.d. uniform r.v. on $[0, 1]$. As a consequence, the cardinality $\#\bar{F}_\varepsilon^{I,c}$ of $\bar{F}_\varepsilon^{I,c}$ is distributed as a Binomial r.v. with parameters $([\frac{c}{\varepsilon}], \varepsilon|I|)$, where $|I|$ refers to the Lebesgue measure of I . Standard arguments yield that $\#\bar{F}_\varepsilon^{I,c}$ converges to a Poisson random variable with parameter $c|I|$.

Next, let $(u^\varepsilon, \sigma^\varepsilon)$ denote the atoms of $(\bar{F}_\varepsilon^{I,c})$. Fix $m \in \mathbb{N}^*$ and for every $\varepsilon > 0$, let us condition on the event $\{\#\bar{F}_\varepsilon^{I,c} = m\}$. (Note that the law of $(T_i; i \geq 0)$ is not affected by the conditioning). This defines a sequence of r.v. $(i_1^\varepsilon, \dots, i_m^\varepsilon)$ of distinct integers in $\{1, \dots, [\frac{c}{\varepsilon}]\}$ such that

$$(u^\varepsilon, \sigma^\varepsilon) = \frac{1}{\varepsilon} \left((V_{i_k^\varepsilon}, T_{i_k^\varepsilon}); k \leq m \right).$$

The statistical description of the comb provided in Proposition 3.1 implies that:

- (i) $(V_{i_k^\varepsilon}; k \leq m)$ are i.i.d. uniform random variables on εI , independent of the sequence $(T_{i_k^\varepsilon}; k \leq m)$.
- (ii) $(i_1^\varepsilon, \dots, i_m^\varepsilon)$ is a uniform sample of size m (with no replacement) of $\{1, \dots, [\frac{c}{\varepsilon}]\}$ independent of the T_i 's.

From (ii), we get:

$$\frac{\varepsilon}{c} (i_1^\varepsilon, \dots, i_m^\varepsilon) \rightarrow (U_1, \dots, U_m) \quad \text{in law}, \quad (5)$$

where U_1, \dots, U_m are i.i.d. uniform r.v. on $[0, 1]$. We now prove that

$$\forall k \leq m, \quad i_k^\varepsilon T_{i_k^\varepsilon} \rightarrow 2 \quad \text{in probability.} \quad (6)$$

In order to see that, we first note that the $i_k T_{i_k}^\varepsilon$'s are exchangeable and to ease the notation, we remove the k -subscript until further notice. For any random variable $X \in L^1$, write $\mathbb{E}_{m,\varepsilon}(X) := \mathbb{E}(X \mid \#\bar{F}_\varepsilon^{I,c} = m)$. Then, using (ii) above:

$$\begin{aligned} \mathbb{E}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon} \mid i^\varepsilon) &= i^\varepsilon \sum_{k \geq i^\varepsilon + 1} \frac{2}{k(k-1)} = \frac{1}{i^\varepsilon} \sum_{k \geq i^\varepsilon + 1} \frac{2i_\varepsilon^2}{k(k-1)} \\ &= 2 \int_1^\infty \frac{dy}{y^2} + R_1(i^\varepsilon) = 2 + R_1(i^\varepsilon), \end{aligned} \quad (7)$$

where R_1 is a function such that $|R_1(x)| \leq K/x$ for some constant K . Averaging of i^ε , we get:

$$\mathbb{E}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon}) = 2 + \sum_{k=1}^{\lfloor \frac{c}{\varepsilon} \rfloor} R_1(k) \mathbb{P}(i^\varepsilon = k) \rightarrow 2, \quad (8)$$

using the fact that $\mathbb{P}(i^\varepsilon = k) = 1/\lfloor \frac{c}{\varepsilon} \rfloor$ and the previous bound on R_1 . Next,

$$\text{Var}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon}) = \mathbb{E}_{m,\varepsilon}(\text{Var}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon} \mid i^\varepsilon)) + \text{Var}_{m,\varepsilon}(\mathbb{E}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon} \mid i^\varepsilon)).$$

First, it is not hard to see from (7) that $\mathbb{E}(i^\varepsilon T_{i^\varepsilon} \mid i^\varepsilon)$ converges to 2 in L^2 , and thus, the second term on the RHS of the latter equality vanishes as $\varepsilon \rightarrow 0$. Let us now deal with the first term

$$\begin{aligned} \text{Var}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon} \mid i^\varepsilon) &= (i^\varepsilon)^2 \sum_{k \geq i^\varepsilon + 1} \left(\frac{2}{k(k-1)} \right)^2 = \frac{1}{i^\varepsilon} \left(\frac{1}{i^\varepsilon} \sum_{k \geq i^\varepsilon + 1} \left(\frac{2(i^\varepsilon)^2}{k(k-1)} \right)^2 \right) \\ &= \frac{1}{i^\varepsilon} \int_1^\infty \frac{4dy}{y^4} + R_2(i^\varepsilon), \end{aligned}$$

where R_2 is a function such that $|R_2(x)| \leq K'/x^2$ for some constant K' . Reasoning as in (8), this shows that the expectation of the RHS of the last equality goes to 0 as $\varepsilon \rightarrow 0$. Altogether, this implies that

$$\text{Var}_{m,\varepsilon}(i^\varepsilon T_{i^\varepsilon}) \rightarrow 0 \text{ as } \varepsilon \rightarrow 0.$$

Together with (8), this completes the proof of (6). Finally, combining (5) with (6) yields the following convergence in law as $\varepsilon \downarrow 0$

$$\left(\frac{1}{\varepsilon} T_{i_1^\varepsilon}, \dots, \frac{1}{\varepsilon} T_{i_m^\varepsilon} \right) \rightarrow \left(\frac{2}{cU_1}, \dots, \frac{2}{cU_m} \right). \quad (9)$$

Let us gather the previous arguments. From the convergence of $\#\bar{F}_\varepsilon^{I,c}$ and (9), it is not hard to deduce that $(\bar{F}_\varepsilon^{I,c})$ is tight. Further, we showed that any sub-sequential limit $\mathcal{F}^{I,c}$ must satisfy the following properties:

1. $\#\mathcal{F}^{I,c}$ is distributed as a Poisson r.v. with parameter $c|I|$.
2. Let $\{(u, \sigma)\}$ denote the atoms of $\mathcal{F}^{I,c}$. Conditional on $\#\mathcal{F}^{I,c}$:
 - (i) (σ) and (u) are independent.
 - (ii) (u) is a sequence of uniform r.v. on I .

(iii) (σ) are i.i.d. r.v. with density

$$\frac{d}{dt} \mathbb{P} \left(\frac{2}{cU_1} \leq t \right) = \frac{d}{dt} \left(1 - \frac{2}{ct} \right) 1_{t > 2/c} = 1_{t > 2/c} \frac{2}{ct^2}.$$

Now these two properties uniquely characterize the law of a PPP with intensity measure

$$2 \int_{x \in I} dx \, 1_{t > \frac{2}{c} \frac{1}{t^2}}.$$

This completes the proof of the convergence of $(\bar{F}_\varepsilon^{I,c})$ and the proof of Proposition 3.4. \blacksquare

3.3 Alternative Proof to Proposition 3.4

The proof of Proposition 3.4 is based on the statistical description of the Kingman comb given in Proposition 3.1. Here, we sketch an alternative proof of Proposition 3.4 relying on totally different techniques, namely Ray-Knight Theorem and a result of Bertoin and Le Gall [BL05] that describes the trajectories of the ancestral lineages in the FV flow of bridges.

Let $x_0 < \dots < x_n$ and assume that ε is small enough such that $\varepsilon x_n < 1$. Define

$$\bar{Y}_i^\varepsilon(t) := \frac{1}{\varepsilon} \phi_{-\varepsilon t, 0}(\varepsilon x_i), \quad i \in \{0, \dots, n\}$$

By Proposition 3.1, $\frac{1}{\varepsilon} d_C(\varepsilon x_{i-1}, \varepsilon x_i)$ coincides with the hitting time at 0 of the process $\bar{Y}_i^\varepsilon(t) - \bar{Y}_{i-1}^\varepsilon(t)$. Thus, it remains to show that this vector of hitting times converges in law to $(\max_{[x_{i-1}, x_i]} \mathcal{C}, i = 1, \dots, n)$, or equivalently, that the components of the vector are asymptotically independent and distributed as $\max_{[0, x_i - x_{i-1}]} \mathcal{C}$ respectively.

Following Theorem 6 in [BL05], the n -point motion $(\bar{Y}_i^\varepsilon, i = 0, \dots, n)$ is a coalescing diffusion whose generator is given by \mathcal{A}^ε :

$$\mathcal{A}^\varepsilon g = \frac{1}{2} \sum_{i,j=0}^n y_{i \wedge j} (1 - \varepsilon y_{i \vee j}) \frac{\partial g}{\partial y_i \partial y_j} + \sum_{i=1}^n \left(\frac{1}{2} - \varepsilon y_i \right) \frac{\partial g}{\partial y_i}, \quad (10)$$

for every $g \in C_0^\infty(\mathbb{R}^{n+1})$. Using standard arguments, one can prove that

$$\left(\frac{1}{\varepsilon} d_C(\varepsilon x_{i-1}, \varepsilon x_i), i = 1, \dots, n \right) \implies (\inf\{t : (\mathcal{Y}_i(t) - \mathcal{Y}_{i-1}(t)) = 0\}, i = 1, \dots, n).$$

where $\{\mathcal{Y}_i\}_{i=0}^n$ is the diffusion with generator \mathcal{A}_0 and initial condition (x_0, \dots, x_n) . The n -dimensional process $\{\mathcal{Y}_i - \mathcal{Y}_{i-1}\}_{i=1}^n$ can be rewritten as MY^0 where M is the $n \times (n+1)$ matrix:

$$M = \begin{pmatrix} -1 & 1 & 0 & 0 & \dots \\ 0 & -1 & 1 & 0 & \dots \\ 0 & 0 & -1 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

Thus, the generator of $\{\mathcal{Y}_i - \mathcal{Y}_{i-1}\}_{i=1}^n$ is given by $\sum_{i=1}^n d_i \frac{\partial}{\partial u_i} + \sum_{i,j=1}^n c_{i,j} \frac{\partial^2}{\partial u_i \partial u_j}$ where

$$\begin{aligned} d &= M \begin{pmatrix} \frac{1}{2} \\ \vdots \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ c &= \frac{1}{2} M [y_{i \wedge j}]_{i,j \leq n} {}^t M. \end{aligned}$$

One can readily check that c is diagonal with the diagonal terms given by $\frac{1}{2}(y_1 - y_0, \dots, y_n - y_{n-1})$. Thus, the processes $(\mathcal{Y}_i - \mathcal{Y}_{i-1})$'s are distributed as independent Feller diffusions (i.e., satisfying Eq (11) below) with respective initial condition $(x_i - x_{i-1})$. The convergence of $S_\varepsilon(C)$ to \mathcal{C} is then a direct consequence of Lemma 3.5 below.

Lemma 3.5. *Let w be a standard Brownian motion. For every $x \in \mathbb{R}$, let σ_x be the hitting time of 0 by the diffusion*

$$dz(t) = \sqrt{z(t)}dw(t), \quad z(0) = x. \quad (11)$$

Then the following identity holds in distribution:

$$\sigma_x = \sup_{[0, x]} \mathcal{C}. \quad (12)$$

Proof. Let $x \geq 0$. Define the local time at x :

$$L^x(t) := \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^t 1_{w(s) \in [x-\varepsilon, x+\varepsilon]} ds, \quad \mathcal{L}^x(t) := \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^t 1_{|w(s)| \in [x-\varepsilon, x+\varepsilon]} ds.$$

and $\tau_y := \inf\{u : L^0(u) > y\}$ to be the inverse local time at 0. In order to prove Lemma 3.5, we will show that the RHS and LHS of (12) are identical in law to $\max_{s \in [0, \tau_{2x}]} |w(s)|$.

We start by showing that the identity holds for the LHS of (12). Define $y(t) = 4z(t)$. Then it is straightforward to check that y is a standard 0-dimensional squared Bessel process (i.e., $dy(t) = \sqrt{2y(t)}dw(t)$) with initial condition $4x$. Furthermore, the hitting time at 0 of y coincides with the one of z . We now construct the process y from the local time of a standard Brownian motion w . By the second Ray-Knight Theorem [K63], the processes

$$(L^t(\tau_{2x}); t \geq 0) \text{ and } (L^{-t}(\tau_{2x}); t \geq 0)$$

are independent 0-dimensional squared Bessel process both starting at $2x$. So their sum $(\mathcal{L}^t(\tau_{2x}); t \geq 0)$ is a 0-dimensional squared Bessel process starting at $4x$, i.e., $(\mathcal{L}^t(\tau_{2x}); t \geq 0)$ has the same distribution as the process y . Finally, since the hitting time at 0 of $(\mathcal{L}^t(\tau_{2x}); t \geq 0)$ coincides with the maximum of $|w|$ on $[0, \tau_{2x}]$, this shows that the LHS of (12) is identical in law to $\max_{s \in [0, \tau_{2x}]} |w(s)|$.

We now show that the same identity holds for the RHS of (12). For every $l \in \mathbb{R}^+$ such that $\tau_{l-} < \tau_l$, define m_l the height of the excursion on the interval $[\tau_{l-}, \tau_l]$ (i.e., $m_l := \sup\{|w(s)| : s \in [\tau_{l-}, \tau_l]\}$). From standard excursion theory (see e.g., Section VI.47 [RW87]), $\sum_l : \tau_{l-} < \tau_l \delta_{l, m_l}$ defines a Poisson Point Process with intensity measure $dl \otimes \frac{dt}{t^2}$. It follows that for every a :

$$\max\{m_l : \tau_{l-} < \tau_l \text{ and } l \leq a\} = \max_{s \in [0, \tau_a]} |w(s)|.$$

Finally since the intensity measure underlying \mathcal{C} is twice the intensity measure of $\sum_l : \tau_{l-} < \tau_l \delta_{l, m_l}$, using standard results on Poisson point processes, it is not hard to show that the RHS of (12) is distributed as $\max_{s \in [0, \tau_{2x}]} |w(s)|$. This completes the proof of Lemma 3.5. \blacksquare

4 Conditional Sampling and Sized-Biased Killed Brownian CPP

Let C be the Kingman comb as defined in Proposition 3.1. For every $\varepsilon > 0$, let us consider the partition \mathcal{P}_ε of $[0, 1]$ induced by the equivalence relation

$$x \sim_{(\varepsilon)} y \iff \phi_{-\varepsilon, 0}(x) = \phi_{-\varepsilon, 0}(y).$$

We let N_ε be the number of equivalence classes (or blocks). To characterize those equivalence classes, let \tilde{C} be the comb that coincides with C on $(0, 1)$ and with $\tilde{C}(0) = \tilde{C}(1) = 1$ and define the sequence $(L_\varepsilon^i)_{i=0}^{N_\varepsilon+1}$ where (L_ε^i) is the ranked enumeration of the set $\{x : \tilde{C}(x) > \varepsilon\}$. It is straightforward to check that the i^{th} block (where blocks are ranked according to their least element) coincides with the interval $(L_\varepsilon^{i-1}, L_\varepsilon^i)$. For $i \leq N_\varepsilon$, we define the comb $\mathcal{B}_\varepsilon^i$ as

$$\mathcal{B}_\varepsilon^i := k_{l_\varepsilon^i} C(\cdot + L_\varepsilon^{i-1}) \text{ with } l_\varepsilon^i := L_\varepsilon^i - L_\varepsilon^{i-1}$$

that can be thought of as the comb encoding the genealogy of the i^{th} block.

This motivates the definition of the comb $M_{n,\varepsilon}$ whose law is characterized as follows. For every bounded measurable $f : \Omega \times (0, \infty) \rightarrow \mathbb{R}$,

$$\mathbb{E}(f(M_{n,\varepsilon}, l_{n,\varepsilon})) := \mathbb{E}\left(\frac{\sum_{i=1}^{N_\varepsilon} (l_\varepsilon^i)^n f(\mathcal{B}_\varepsilon^i, l_\varepsilon^i)}{\sum_{i=1}^{N_\varepsilon} (l_\varepsilon^i)^n}\right). \quad (13)$$

In words, conditioned on a realization of the population, we first generate n independent uniform random variables on $[0, 1]$, independent of the flow of bridges. If we condition those random variables to fall into the same block, $l_{n,\varepsilon}$ is defined as the size of the chosen block and $M_{n,\varepsilon}$ is the comb encoding the genealogy of the (averaged) block. The aim of this section is to prove the following result.

Theorem 4.1. *Let \mathcal{M} be a CPP with intensity measure $\nu(dl)1_{l \leq 1}$ and \mathcal{L}_n be an independent Gamma distributed random variable with parameter $(n+1, 2)$. Then the following convergence*

$$\left(S_\varepsilon(M_{n,\varepsilon}), \frac{1}{\varepsilon} l_{n,\varepsilon}\right) \implies (k_{\mathcal{L}_n}(\mathcal{M}), \mathcal{L}_n)$$

holds weakly in law as $\varepsilon \rightarrow 0$.

Remark 4.2. *As already mentioned in the introduction, the (properly rescaled) coalescent point process associated to a critical birth-death process conditioned on surviving up to time N converges to $k_{\mathcal{L}_0}(\mathcal{M})$ as $N \rightarrow \infty$ [P04]. As a consequence, the previous result states that after rescaling, the limit of $(M_{n,\varepsilon})$ is described in terms a ‘ n -size-biased critical birth-death process’ (i.e., biased by the n -th power of its size) conditioned on survival up to a large time.*

4.1 The n^{th} moment of the length of a uniformly chosen block

In order to prove Theorem 4.1, we will need to first establish some technical results. Define

$$X_{n,\varepsilon} = \frac{1}{N_\varepsilon} \sum_{i=1}^{N_\varepsilon} \left(\frac{2}{\varepsilon} l_\varepsilon^i\right)^n.$$

The aim of this section is to prove the following proposition.

Proposition 4.3. *As $\varepsilon \rightarrow 0$, the sequence $(X_{n,\varepsilon})$ converges to $n!$ in L^2 .*

Before proceeding with the proof, we first deduce an easy corollary of Proposition 4.3.

Corollary 4.4. *The collection of r.v. $\{(\frac{2}{\varepsilon}l_1^\varepsilon)^n\}_{\varepsilon>0}$ is uniformly integrable.*

Proof. Let $V_\varepsilon = (\frac{2}{\varepsilon}l_1^\varepsilon)^n$. Applying Cauchy–Schwartz inequality:

$$\begin{aligned} \mathbb{E}(V_\varepsilon 1_{V_\varepsilon > M})^2 &\leq \mathbb{E}(V_\varepsilon^2) \mathbb{P}(V_\varepsilon > M) \\ &= \mathbb{E} \left(\mathbb{E} \left(\left(\frac{2}{\varepsilon} l_1^\varepsilon \right)^{2n} \middle| N_\varepsilon \right) \right) \mathbb{P}(V_\varepsilon > M) \\ &= \mathbb{E} \left(\mathbb{E} \left(\frac{1}{N_\varepsilon} \sum_{i=1}^{N_\varepsilon} \left(\frac{2}{\varepsilon} l_i^\varepsilon \right)^{2n} \middle| N_\varepsilon \right) \right) \mathbb{P}(V_\varepsilon > M) \\ &= \mathbb{E}(X_{2n,\varepsilon}) \mathbb{P}(V_\varepsilon > M), \end{aligned}$$

where the second equality follows from the fact that conditioned on N_ε , the random variables $\{l_i^\varepsilon\}_{i=1}^{N_\varepsilon}$ are exchangeable. On the one hand, Proposition 4.3 implies that

$$\sup_{\varepsilon>0} \mathbb{E}(X_{2n,\varepsilon}) < \infty.$$

On the other hand, Proposition 3.4 implies that V_ε converges in law to $(l_1(\mathcal{C}))^n$, and thus $\{V_\varepsilon\}_{\varepsilon>0}$ is tight. This ends the proof of the lemma. \blacksquare

We now proceed with the proof of Proposition 4.3. We first need to define the r.v. $Z_{n,\varepsilon}$. Let us consider a sequence $\{e_k\}_{k \geq 1}$ of independent random variables where e_k is an exponential random variable with parameter $\frac{k(k-1)}{2}$ and an independent sequence $\{\xi_k\}_{k \geq 1}$ of i.i.d exponential random variables with parameter 1. Now set

$$Z_{n,\varepsilon} := \left(\frac{1}{\bar{N}_\varepsilon} \sum_{k=1}^{\bar{N}_\varepsilon} \left(\frac{\xi_k}{S_{\bar{N}_\varepsilon}} \frac{2}{\varepsilon} \right)^n \right), \quad (14)$$

where $\bar{N}_\varepsilon := \inf\{n : \sum_{k \geq n} e_k < \varepsilon\}$, and $S_n = \sum_{k=1}^n \xi_k$.

Lemma 4.5. *For all n and ε , the r.v. $Z_{n,\varepsilon}$ and $X_{n,\varepsilon}$ are equally distributed.*

Proof. For the standard Fleming-Viot flow of bridges, the number of blocks N_ε is distributed as the value at time ε of a pure death process descending from ∞ with rate $k(k-1)/2$ at level k , and further, conditioned on N_ε , the vector $\{l_i^\varepsilon\}_{i=1}^{N_\varepsilon}$ is distributed as a N_ε -dimensional Dirichlet random variable with parameter $(1, \dots, 1)$. Finally, it is well known that the n -dimensional Dirichlet random variable with parameter $(1, \dots, 1)$ is distributed as

$$(\xi_1/S_n, \dots, \xi_n/S_n),$$

which completes the proof of the lemma. \blacksquare

Lemma 4.6. *Let $\gamma > 0$ and define*

$$N_{\varepsilon,\gamma}^- = \left\lfloor (1-\gamma) \frac{2}{\varepsilon} \right\rfloor, \text{ and } N_{\varepsilon,\gamma}^+ = \left\lceil (1+\gamma) \frac{2}{\varepsilon} \right\rceil.$$

Then for any $k \in \mathbb{N}$, $\mathbb{P}(\bar{N}_\varepsilon \notin [N_{\varepsilon,\gamma}^-, N_{\varepsilon,\gamma}^+]) / \varepsilon^k \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Proof. If $R_N = \sum_{k \geq N} e_k$, then

$$\mathbb{E}(\exp(-\lambda N^2 R_N)) = \exp \left(\sum_{k \geq N} \log \left(\frac{\frac{k}{2N}(\frac{k}{N} - \frac{1}{N})}{\lambda + \frac{k}{2N}(\frac{k}{N} - \frac{1}{N})} \right) \right)$$

For every $\lambda > -1/2$, the function $u \rightarrow \ln \left(\frac{\frac{u^2}{2}}{\frac{u^2}{2} + \lambda} \right)$ is integrable on $[1, \infty)$. Thus, as $N \rightarrow \infty$, it is not hard to show the following convergence

$$\frac{1}{N} \sum_{k \geq N} \log \left(\frac{\frac{k}{2N}(\frac{k}{N} - \frac{1}{N})}{\lambda + \frac{k}{2N}(\frac{k}{N} - \frac{1}{N})} \right) \rightarrow \int_1^\infty \ln \left(\frac{\frac{u^2}{2}}{\frac{u^2}{2} + \lambda} \right) du, \quad (15)$$

or equivalently,

$$\mathbb{E}(\exp(-\lambda N^2 R_N)) = \exp \left(-N \int_1^\infty \ln \left(\frac{\frac{u^2}{2}}{\frac{u^2}{2} + \lambda} \right) du \right) \gamma_N(\lambda). \quad (16)$$

with $\gamma_N(\lambda) \rightarrow 1$ as $N \rightarrow \infty$. From Chebyshev's inequality this implies:

$$\begin{aligned} \mathbb{P}(NR_N < 2(1 - \gamma)) &\leq \exp(\lambda N 2(1 - \gamma)) \mathbb{E}(\exp(-\lambda N^2 R_N)) \\ &= \exp(N(\lambda 2(1 - \gamma) - f(\lambda))) \gamma_N(\lambda) \\ \text{with } f(\lambda) &= \int_1^\infty \ln \left(\frac{\frac{u^2}{2}}{\frac{u^2}{2} + \lambda} \right) du. \end{aligned}$$

Since $f(0) = 0$ and $f'(0) = 2$, we can choose $\lambda > 0$ small enough such that $\lambda(2 - \gamma) - f(\lambda) < 0$. By the same token, for every $\lambda \in (-1/2, 0)$ and $N \in \mathbb{N}$:

$$\begin{aligned} \mathbb{P}(NR_N \geq 2(1 + \gamma)) &< \exp(\lambda N 2(1 + \gamma)) \mathbb{E}(\exp(-\lambda N^2 R_N)) \\ &= \exp(N(\lambda 2(1 + \gamma) - f(\lambda))) \gamma_N(\lambda) \end{aligned}$$

and again, we can choose $\lambda < 0$ small enough such that $\lambda(2 + \gamma) - f(\lambda) < 0$. To complete the proof, we combine the previous large deviation estimates with the observation

$$\{\bar{N}_\varepsilon < N_{\varepsilon, \gamma}^+\} = \{R_{N_{\varepsilon, \gamma}^+} \geq \varepsilon\} = \{N_{\varepsilon, \gamma}^+ R_{N_{\varepsilon, \gamma}^+} \geq N_{\varepsilon, \gamma}^+ \varepsilon\}$$

with $N_{\varepsilon, \gamma}^+ \varepsilon = 2(1 + \gamma)$ and

$$\{\bar{N}_\varepsilon \geq N_{\varepsilon, \gamma}^-\} = \{N_{\varepsilon, \gamma}^- R_{N_{\varepsilon, \gamma}^-} \leq N_{\varepsilon, \gamma}^- \varepsilon\}$$

with $N_{\varepsilon, \gamma}^- \varepsilon = 2(1 - \gamma)$. Combining this with the previous inequalities shows that $\mathbb{P}(\bar{N}_\varepsilon \notin [N_{\varepsilon, \gamma}^-, N_{\varepsilon, \gamma}^+])$ goes exponentially fast to 0 in ε as $\varepsilon \rightarrow 0$. This completes the proof of Lemma 4.6. \blacksquare

Proof of Proposition 4.3. From Lemma 4.5, it is enough to show that $Z_{n, \varepsilon}$ converges to $n!$ in L^2 . To ease the notation, we drop the dependence in n and write $Z_\varepsilon \equiv Z_{n, \varepsilon}$. Let us now introduce two auxiliary variables:

$$Z_\varepsilon^+ = \left(\frac{1}{N_{\varepsilon, \gamma}^+} \sum_{k=1}^{N_{\varepsilon, \gamma}^+} \left(\frac{\xi_k}{S_{N_{\varepsilon, \gamma}^+}} \frac{2}{\varepsilon} \right)^n \right), \quad Z_\varepsilon^- = \left(\frac{1}{N_{\varepsilon, \gamma}^-} \sum_{k=1}^{N_{\varepsilon, \gamma}^-} \left(\frac{\xi_k}{S_{N_{\varepsilon, \gamma}^-}} \frac{2}{\varepsilon} \right)^n \right)$$

Since $Z_\varepsilon < \left(\frac{2}{\varepsilon}\right)^n$ and $Z_\varepsilon^- \leq Z_\varepsilon \leq Z_\varepsilon^+$ on $\{\bar{N}_\varepsilon \in [N_{\varepsilon,\gamma}^-, N_{\varepsilon,\gamma}^+]\}$, for every $\gamma > 0$:

$$\mathbb{E}(Z_\varepsilon - n!)^2 < \left(n! + \left(\frac{2}{\varepsilon}\right)^n\right)^2 \mathbb{P}(\bar{N}_\varepsilon \notin [N_{\varepsilon,\gamma}^-, N_{\varepsilon,\gamma}^+]) + \mathbb{E}(Z_{\varepsilon,\gamma}^+ - n!)^2 + \mathbb{E}(Z_{\varepsilon,\gamma}^- - n!)^2.$$

Lemma 4.6 implies that the first term vanishes as $\varepsilon \rightarrow 0$ and it remains to show that

$$\lim_{\gamma \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \mathbb{E}(Z_{\varepsilon,\gamma}^+ - n!)^2, \mathbb{E}(Z_{\varepsilon,\gamma}^- - n!)^2 = 0$$

We only show the first limit. The second limit can be shown along the same lines. We can rewrite

$$\begin{aligned} Z_{\varepsilon,\gamma}^+ &= \frac{N_{\varepsilon,\gamma}^+}{N_{\varepsilon,\gamma}^-} (1-\gamma)^n \frac{1}{N_{\varepsilon,\gamma}^+} \sum_{k=1}^{N_{\varepsilon,\gamma}^+} \left(\frac{\xi_k}{S_{N_{\varepsilon,\gamma}^-}} \frac{2}{\varepsilon(1-\gamma)} \right)^n \\ &= c_{\varepsilon,\gamma} \underbrace{\frac{1}{N_{\varepsilon,\gamma}^+} \sum_{k=1}^{N_{\varepsilon,\gamma}^+} (\xi_k)^n}_{:= A_{N_{\varepsilon,\gamma}^+}} \underbrace{\left(\frac{N_{\varepsilon,\gamma}^-}{S_{N_{\varepsilon,\gamma}^-}} \right)^n}_{:= B_{N_{\varepsilon,\gamma}^-}} \end{aligned}$$

with

$$c_{\varepsilon,\gamma} = \frac{N_{\varepsilon,\gamma}^+}{N_{\varepsilon,\gamma}^-} (1-\gamma)^n \left(\frac{2N_{\varepsilon,\gamma}^-}{\varepsilon} \right)^n \text{ so that } \lim_{\gamma \rightarrow 0} \lim_{\varepsilon \rightarrow 0} c_{\varepsilon,\gamma} = 1.$$

Next, writing $A = \mathbb{E}(\xi^n) = n!$ and $B = \left(\frac{1}{\mathbb{E}(\xi)}\right)^n = 1$, we have

$$\begin{aligned} A_{N_{\varepsilon,\gamma}^+} B_{N_{\varepsilon,\gamma}^-} - AB &= B(A_{N_{\varepsilon,\gamma}^+} - A) + A(B_{N_{\varepsilon,\gamma}^-} - B) + (A_{N_{\varepsilon,\gamma}^+} - A)(B_{N_{\varepsilon,\gamma}^-} - B) \\ &= B(A_{N_{\varepsilon,\gamma}^+} - A) + A_{N_{\varepsilon,\gamma}^+} B_{N_{\varepsilon,\gamma}^-} \left(1 - \frac{B}{B_{N_{\varepsilon,\gamma}^-}}\right) + B_{N_{\varepsilon,\gamma}^-} (A_{N_{\varepsilon,\gamma}^+} - A) \left(1 - \frac{B}{B_{N_{\varepsilon,\gamma}^-}}\right). \end{aligned}$$

and in order to prove Proposition 4.3, it remains to show that the RHS of the equality goes to 0 in L^2 as ε goes to 0. By applying Cauchy–Schwartz inequality several times, it is easy to show that the RHS converges to 0 in L^2 if for all $p \in \mathbb{N}$,

$$(A_N - A) \rightarrow 0 \text{ and } \left(\frac{B}{B_N} - 1\right) \rightarrow 0 \text{ as } N \rightarrow \infty \text{ in } L^p \quad (17)$$

$$\text{and } \sup_N \mathbb{E}(A_N^p) < \infty \text{ and } \sup_N \mathbb{E}(B_N^p) < \infty. \quad (18)$$

Condition (17) is a law of large number and can easily be checked using the fact that the ξ_i 's are i.i.d exponential random variables. The first condition of (18) directly follows from the first part of (17). For the second assertion of (18), we first note that S_N is distributed as a Gamma distribution with parameter $(N, 1)$, and thus:

$$\mathbb{E}(B_N^p) = N^{np} \frac{1}{\Gamma(N)} \int_0^\infty x^{N-1-np} e^{-x} dx = N^{np} \frac{(N-np-1)!}{(N-1)!} \rightarrow 1$$

as $N \rightarrow \infty$. This ends the proof of Proposition 4.3. ■

4.2 Proof of Theorem 4.1

We now proceed with the proof of Theorem 4.1. Let $0 < x_0 < x_1 < \dots < x_{n-1}$ and let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be an arbitrary bounded and continuous function. Define $f : \Omega \times (0, \infty) \rightarrow \mathbb{R}$ as

$$f(\omega, x) := g(d_\omega(x_0, x_1), \dots, d_\omega(x_{n-2}, x_{n-1}), x).$$

We aim at showing that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}(f(S_\varepsilon(M_{n,\varepsilon}), l_{n,\varepsilon}/\varepsilon)) = \mathbb{E}(f(k_{\mathcal{L}_n}(\mathcal{M}), \mathcal{L}_n)), \quad (19)$$

where $(\mathcal{M}, \mathcal{L}_n)$ are defined as in Theorem 4.1. Set $S_\varepsilon(\mathcal{B}_\varepsilon^i, l_\varepsilon^i) = (S_\varepsilon(\mathcal{B}_\varepsilon^i), \frac{1}{\varepsilon} l_\varepsilon^i)$ and

$$x_\varepsilon = \sum_{i=1}^{N_\varepsilon} f \circ S_\varepsilon(\mathcal{B}_\varepsilon^i, l_\varepsilon^i) (l_\varepsilon^i)^\varepsilon / \sum_{i=1}^{N_\varepsilon} (l_\varepsilon^i)^\varepsilon, \quad y_\varepsilon = \frac{1}{n!} \frac{1}{N_\varepsilon} \sum_{i=1}^{N_\varepsilon} f \circ S_\varepsilon(\mathcal{B}_\varepsilon^i, l_\varepsilon^i) \left(\frac{2}{\varepsilon} l_\varepsilon^i\right)^\varepsilon$$

Straightforward manipulation yields that for all $\varepsilon > 0$,

$$|x_\varepsilon - y_\varepsilon| \leq \|g\|_\infty \left| 1 - \frac{X_{n,\varepsilon}}{n!} \right|.$$

where $X_{n,\varepsilon}$ is defined as in Section 4.1. From Proposition 4.3, the RHS goes to 0 in L^1 as $\varepsilon \rightarrow 0$. Now by definition

$$\mathbb{E}(f(S_\varepsilon(M_{n,\varepsilon}), l_{n,\varepsilon}/\varepsilon)) = \mathbb{E}(x_\varepsilon),$$

and thus

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}(f(S_\varepsilon(M_{n,\varepsilon}), l_{n,\varepsilon}/\varepsilon)) = \frac{1}{n!} \mathbb{E} \left(\frac{1}{N_\varepsilon} \sum_{i=1}^{N_\varepsilon} f \circ S_\varepsilon(\mathcal{B}_\varepsilon^i, l_\varepsilon^i) \left(\frac{2}{\varepsilon} l_\varepsilon^i\right)^\varepsilon \right) = 0.$$

Next, using exchangeability, we get that

$$\begin{aligned} \mathbb{E} \left(\frac{1}{N_\varepsilon} \sum_{i=1}^{N_\varepsilon} f \circ S_\varepsilon(\mathcal{B}_\varepsilon^i, l_\varepsilon^i) \left(\frac{2}{\varepsilon} l_\varepsilon^i\right)^\varepsilon \right) &= \mathbb{E} \left(\mathbb{E} \left(\frac{1}{N_\varepsilon} \sum_{i=1}^{N_\varepsilon} f \circ S_\varepsilon(\mathcal{B}_\varepsilon^i, l_\varepsilon^i) \left(\frac{2}{\varepsilon} l_\varepsilon^i\right)^\varepsilon \middle| N_\varepsilon \right) \right) \\ &= \mathbb{E} \left(\mathbb{E} \left(f \circ S_\varepsilon(\mathcal{B}_\varepsilon^1, l_\varepsilon^1) \left(\frac{2}{\varepsilon} l_\varepsilon^1\right)^\varepsilon \middle| N_\varepsilon \right) \right) \\ &= \mathbb{E} \left(f \circ S_\varepsilon(\mathcal{B}_\varepsilon^1, l_\varepsilon^1) \left(\frac{2}{\varepsilon} l_\varepsilon^1\right)^\varepsilon \right). \end{aligned}$$

By Proposition 3.4, $S_\varepsilon(\mathcal{B}_\varepsilon^1, l_\varepsilon^1)$ converges weakly to $(\bar{\mathcal{C}}, l_1(\mathcal{C}))$, where $\bar{\mathcal{C}} = k_{l_1} \mathcal{C}$. Using the uniform integrability of $((\frac{2}{\varepsilon} l_\varepsilon^1)^\varepsilon)_{\varepsilon > 0}$ (by Corollary 4.4), this yields

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}(f(S_\varepsilon(M_{n,\varepsilon}), l_{n,\varepsilon}/\varepsilon)) = \frac{2^n}{n!} \mathbb{E}(f(\bar{\mathcal{C}}, l_1(\mathcal{C})) l_1(\mathcal{C})^n). \quad (20)$$

To complete the proof of (19) (and thus of Theorem 4.1), we first note that $l_1(\mathcal{C})$ is an exponential random variable with parameter 2, since

$$\mathbb{P}(l_1(\mathcal{C}) \geq x) = \exp \left(-2 \int_{l \in (0, x)} \int_{u \in [1, \infty)} \frac{du \, dl}{u^2} \right) = \exp(-2x),$$

Second, note that $(\bar{\mathcal{C}}, l_1(\mathcal{C}))$ is identical in law to the pair $(\bar{\mathcal{C}}', l')$ where $\bar{\mathcal{C}}'$ is obtained by killing at the exponential random variable l' an independent CPP \mathcal{M} with intensity measure $\nu(dl)1_{l < 1}$. In other words

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}(f(S_\varepsilon(M_{n,\varepsilon}), l_{n,\varepsilon}/\varepsilon)) = \frac{2^n}{n!} \int_0^\infty dx 2e^{-2x} \mathbb{E}(f(k_x(\mathcal{M}), x)) x^n,$$

which ends the proof of Theorem 4.1.

5 Genealogies Associated with Two Conditional Samplings

5.1 Quenched Conditional Sampling

We now consider the genealogy of the n uniformly sampled individuals after quenched conditional sampling as defined in the previous section (see Equation (13)). For a given realization of $M_{n,\varepsilon}$, let U_0, \dots, U_{n-1} be n independent uniform random variables on the interval $[0, l_{n,\varepsilon}]$. Let $(U_{(0)}, \dots, U_{(n-1)})$ be the vector (U_0, \dots, U_{n-1}) reordered from least to greatest and define

$$H_i^\varepsilon := d_{M_{n,\varepsilon}}(U_{(i-1)}, U_{(i)}) \quad i \in \{1, \dots, n-1\}$$

the coalescence times of our sample.

Theorem 5.1. *As $\varepsilon \rightarrow 0$, the coalescence times $(\varepsilon^{-1}H_i^\varepsilon, i \in \{1, \dots, n-1\})$ converge to $n-1$ i.i.d. uniformly distributed random variables on $(0, 1)$.*

In order to show Theorem 5.1, we define the \mathcal{H}_i 's analogously to H_i^ε , but with respect to $k_{\mathcal{L}_n}(\mathcal{M})$, that is

$$\mathcal{H}_i := d_{\mathcal{M}}(\mathcal{U}_{(i-1)}, \mathcal{U}_{(i)}),$$

where the $\{\mathcal{U}_{(i)}\}_{i=0}^{n-1}$ form the reordering of n independent uniform random variables on the interval $[0, \mathcal{L}_n]$. We decompose the proof of Theorem 5.1 into two steps. We first show that $\{H_i^\varepsilon\}$ converges to $\{\mathcal{H}_i\}$ (see Lemma 5.2). We then characterize the distribution of the \mathcal{H}_i 's (see Lemma 5.3).

Lemma 5.2. *As $\varepsilon \rightarrow 0$, we have the convergence in distribution of $(\varepsilon^{-1}H_i^\varepsilon, i \in \{1, \dots, n-1\})$ to $(\mathcal{H}_i, i \in \{1, \dots, n-1\})$.*

Proof. Let f be a bounded continuous function from \mathbb{R}^{n-1} to \mathbb{R} and let $x_0 < x_1 < \dots < x_{n-1}$. Define

$$\begin{aligned} g_\varepsilon(x_0, x_1, \dots, x_{n-1}) &= \mathbb{E} \left(f(d_{S_\varepsilon(M_{n,\varepsilon})}(x_0, x_1), \dots, d_{S_\varepsilon(M_{n,\varepsilon})}(x_{n-2}, x_{n-1})) 1_{\varepsilon x_{n-1} < l_{n,\varepsilon}} \frac{\varepsilon^n}{l_{n,\varepsilon}^n} \right) \\ g_0(x_0, x_1, \dots, x_{n-1}) &= \mathbb{E} \left(f(d_{\mathcal{M}}(x_0, x_1), \dots, d_{\mathcal{M}}(x_{n-2}, x_{n-1})) 1_{x_{n-1} < \mathcal{L}_n} \frac{1}{\mathcal{L}_n^n} \right). \end{aligned}$$

It is not hard to see that

$$\mathbb{E}(f(\varepsilon^{-1}H_1^\varepsilon, \dots, \varepsilon^{-1}H_{n-1}^\varepsilon)) = n! \int_{\Delta_n} g_\varepsilon(\vec{x}) d\lambda(\vec{x})$$

where λ is the Lebesgue measure on \mathbb{R}^n and $\Delta_n = \{\vec{x} \in \mathbb{R}^n : x_0 < x_1 < \dots < x_{n-1}\}$. Further, an analogous relation holds for the limiting object. Thus we need to show that

$$\lim_{\varepsilon \rightarrow 0} \int_{\Delta_n} g_\varepsilon(\vec{x}) d\lambda(\vec{x}) = \int_{\Delta_n} g_0(\vec{x}) d\lambda(\vec{x}). \quad (21)$$

Since $\mathbb{P}(\mathcal{L}_n = x) = 0$ for every $x \in \mathbb{R}$, Theorem 4.1 implies that g_ε converges almost surely to g_0 on $\{0 < x_1 < \dots < x_{n-1}\}$, and thus, it remains to justify the limit-integral interchange in (21). For every $\delta < A \in (0, \infty)$:

$$\begin{aligned} \left| \int_{\Delta_n} (g_\varepsilon(\vec{x}) - g_0(\vec{x})) d\lambda(\vec{x}) \right| &< \int_{\Delta_n \cap \{x_{n-1} > A\}} |g_\varepsilon(\vec{x})| d\lambda(\vec{x}) + \int_{\Delta_n \cap \{x_{n-1} > A\}} |g_0(\vec{x})| d\lambda(\vec{x}) \\ &+ \int_{\Delta_n \cap \{x_{n-1} < \delta\}} |g_\varepsilon(\vec{x})| d\lambda(\vec{x}) + \int_{\Delta_n \cap \{x_{n-1} < \delta\}} |g_0(\vec{x})| d\lambda(\vec{x}) \\ &+ \left| \int_{\Delta_n \cap \{x_{n-1} \in [\delta, A]\}} (g_\varepsilon(\vec{x}) - g_0(\vec{x})) d\lambda(\vec{x}) \right|. \end{aligned} \quad (22)$$

The last term goes to 0 by combining the bounded convergence theorem and Theorem 4.1. Next, let M be such that $\|f\|_\infty < M$

$$\begin{aligned} \int_{\Delta_n \cap \{x_{n-1} > A\}} |g_\varepsilon(\vec{x})| d\lambda(\vec{x}) &< M \int_{\Delta_n \cap \{x_{n-1} > A\}} \mathbb{E} \left(1_{\varepsilon x_{n-1} < l_{n,\varepsilon} \frac{\varepsilon^n}{l_{n,\varepsilon}^n}} \right) d\lambda(\vec{x}) \\ &= M \mathbb{E} \left(\int_{\Delta_n} 1_{A < x_{n-1} < l_{n,\varepsilon}/\varepsilon} \frac{\varepsilon^n}{l_{n,\varepsilon}^n} d\lambda(\vec{x}) \right) \\ &= \frac{M}{n!} \mathbb{E}(1_{l_{n,\varepsilon}/\varepsilon > A} \left(1 - \left(\frac{\varepsilon A}{l_{n,\varepsilon}} \right)^n \right)) < \frac{M}{n!} \mathbb{P}(l_{n,\varepsilon}/\varepsilon > A) \end{aligned}$$

By a similar computation, we get

$$\int_{\Delta_n \cap \{x_{n-1} > A\}} |g_0(\vec{x})| d\lambda(\vec{x}) < \frac{M}{n!} \mathbb{P}(\mathcal{L}_n > A)$$

Next, we control the second line of (22).

$$\begin{aligned} \int_{\Delta_n \cap \{x_{n-1} < \delta\}} |g_\varepsilon(\vec{x})| d\lambda(\vec{x}) &\leq M \mathbb{E} \left(\int_{\Delta_n \cap \{x_{n-1} < \delta\}} 1_{x_{n-1} < l_{n,\varepsilon}/\varepsilon} \frac{\varepsilon^n}{l_{n,\varepsilon}^n} d\lambda(\vec{x}), l_{n,\varepsilon}/\varepsilon < \delta \right) \\ &+ M \mathbb{E} \left(\int_{\Delta_n \cap \{x_{n-1} < \delta\}} \frac{\varepsilon^n}{l_{n,\varepsilon}^n} d\lambda(\vec{x}), l_{n,\varepsilon}/\varepsilon > \delta \right) \\ &= \frac{M}{n!} \left(\mathbb{P}(l_{n,\varepsilon}/\varepsilon < \delta) + \delta^n \mathbb{E} \left(\frac{\varepsilon^n}{l_{n,\varepsilon}^n}, l_{n,\varepsilon}/\varepsilon > \delta \right) \right). \end{aligned}$$

By a similar computation, we get

$$\int_{\Delta_n \cap \{x_{n-1} < \delta\}} |g_0(\vec{x})| d\lambda(\vec{x}) \leq \frac{M}{n!} \left(\mathbb{P}(\mathcal{L}_n < \delta) + \delta^n \mathbb{E} \left(\frac{1}{\mathcal{L}_n^n}, \mathcal{L}_n > \delta \right) \right).$$

Collecting the previous bounds, using the fact that $l_{n,\varepsilon}/\varepsilon \implies \mathcal{L}_n$ and the bounded convergence theorem yield that for every $\delta < A \in (0, \infty)$

$$\limsup_{\varepsilon \downarrow 0} \left| \int_{\Delta_n} (g_\varepsilon(\vec{x}) - g_0(\vec{x})) d\lambda(\vec{x}) \right| \leq \frac{2M}{n!} \left(\mathbb{P}(\mathcal{L}_n \geq A) + \mathbb{P}(\mathcal{L}_n < \delta) + \delta^n \mathbb{E} \left(\frac{1}{\mathcal{L}_n^n}, \mathcal{L}_n > \delta \right) \right) \quad (23)$$

Since \mathcal{L}_n is distributed as a Gamma random variable with parameter $(n+1, 2)$, we have

$$\begin{aligned} \limsup_{\delta \rightarrow 0} \delta^n \mathbb{E}\left(\frac{1}{\mathcal{L}_n^n}, \mathcal{L}_n > \delta\right) &= \frac{2^n}{n!} \limsup_{\delta \rightarrow 0} \delta^n \int_{\delta}^{\infty} e^{-2x} dx \\ &= 0. \end{aligned}$$

Taking the limit $\delta \rightarrow 0$ and $A \rightarrow \infty$ in (23) yields the desired result. \blacksquare

We now characterize the distribution of the vector $(\mathcal{H}_i, i \in \{1, \dots, n-1\})$.

Lemma 5.3. *The r.v. $(\mathcal{H}_i, i \in \{1, \dots, n-1\})$ are i.i.d. uniformly distributed on $[0, 1]$.*

Proof. Recall that \mathcal{L}_n is a Gamma r.v. with parameter $(n+1, 2)$ and that conditional on \mathcal{L}_n , the $\{\mathcal{U}_{(i)}\}_{i=0}^{n-1}$ form the reordering of n independent uniform random variables $\{\mathcal{U}_i\}_{i=0}^{n-1}$ on the interval $[0, \mathcal{L}_n]$. So the $(n+1)$ -tuple $(\mathcal{U}_0, \dots, \mathcal{U}_{n-1}, \mathcal{L}_n)$ has measure on \mathbb{R}^{n+1}

$$\frac{2^{n+1}}{n!} \exp(-2L) L^n dL \times \frac{du_0}{L} 1_{0 < u_0 < L} \times \dots \times \frac{du_{n-1}}{L} 1_{0 < u_{n-1} < L},$$

that is, $(\mathcal{U}_0, \dots, \mathcal{U}_{n-1}, \mathcal{L}_n)$ has a density equal to $\frac{2^{n+1}}{n!} e^{-2L}$ on the following subset of \mathbb{R}_+^{n+1}

$$A := \{(u_0, \dots, u_{n-1}, L) : \forall i, 0 < u_i < L\}.$$

Now there are $n!$ pairwise disjoint $\{A_i\}_{i=1}^{n!}$ open subsets of A corresponding to the $n!$ possible orderings of (u_0, \dots, u_{n-1}) and whose union coincides with A λ -a.e. In addition, on each A_i , the map

$$\begin{aligned} A_i &\rightarrow \mathbb{R}_+^{n+1} \\ (u_0, \dots, u_{n-1}, L) &\mapsto (u_{(0)}, u_{(1)} - u_{(0)}, \dots, u_{(n-1)} - u_{(n-2)}, L - u_{(n-1)}) \end{aligned}$$

is a C^1 -diffeomorphism with jacobian equal to 1 in absolute value. As a consequence, the $(n+1)$ -tuple $(\mathcal{U}_{(0)}, \mathcal{U}_{(1)} - \mathcal{U}_{(0)}, \dots, \mathcal{U}_{(n-1)} - \mathcal{U}_{(n-2)}, \mathcal{U}_{(n)} - \mathcal{U}_{(n-1)})$, where we have set $\mathcal{U}_{(n)} := \mathcal{L}_n$ has a density on \mathbb{R}_+^{n+1} equal to

$$2^{n+1} e^{-2L} = \prod_{i=0}^n 2 e^{-2(u_{(i)} - u_{(i-1)})},$$

(denoting $u_{(-1)} = 0$), which proves that these r.v. are i.i.d. exponential with parameter 2. Now observe that for any $\sigma \in [0, 1]$ and $0 < a < b$,

$$\mathbb{P}(\sup\{\mathcal{M}(x), x \in [a, b]\} \leq \sigma) = \exp\left(-(b-a) \int_{\sigma}^1 \frac{2ds}{s^2}\right) = \exp(-2(b-a)(\sigma^{-1} - 1)).$$

From the fact that \mathcal{M} is a coalescent point process, for every $(n-1)$ -tuple $\{\sigma_i\}_{i=1}^{n-1}$ of $[0, 1]$,

$$\begin{aligned} \mathbb{P}(\sup\{\mathcal{M}(x), x \in [\mathcal{U}_{(i-1)}, \mathcal{U}_{(i)}]\} \leq \sigma_i, \forall i) &= \mathbb{E}\left(\prod_{i=1}^{n-1} \exp(-2(\mathcal{U}_{(i)} - \mathcal{U}_{(i-1)})(\sigma_i^{-1} - 1))\right) \\ &= \prod_{i=1}^{n-1} \int_0^{\infty} dx 2e^{-2x} \exp(-2x(\sigma_i^{-1} - 1)) \\ &= \prod_{i=1}^{n-1} \sigma_i, \end{aligned}$$

where we have used that $\{\mathcal{U}_{(i)} - \mathcal{U}_{(i-1)}\}_{i=1}^{n-1}$ is a vector of i.i.d. exponential random variables with parameter 2. \blacksquare

5.2 Averaged Conditional Sampling

We now consider the simpler conditional sampling scheme, namely we consider a uniform sample of n individuals and condition its genealogy to coalesce at a depth smaller than ε . In other words, we consider $T_n < T_{n-1} < \dots < T_2$ the successive jump times in the n -Kingman coalescent and we define $(T_i^\varepsilon, i \in \{2, \dots, n\})$ as the $(n-1)$ -tuple $(\varepsilon^{-1}T_i, i \in \{2, \dots, n\})$ conditional on the event $\{T_2 < \varepsilon\}$.

Theorem 5.4. *As $\varepsilon \rightarrow 0$, the ranked coalescence times $(T_i^\varepsilon, i \in \{2, \dots, n\})$ converge to the order statistics of $n-1$ i.i.d. uniformly distributed random variables on $(0, 1)$.*

Proof. For any $i \geq 2$ set $a_i := i(i-1)/2$ and let $0 < t_n < t_{n-1} < \dots < t_2 < 1$. Then setting $T_{n+1} := 0$, we get

$$\begin{aligned} \mathbb{P}(\varepsilon^{-1}T_i \in dt_i, 2 \leq i \leq n) &= \mathbb{P}(\varepsilon^{-1}(T_i - T_{i+1}) \in dt_i - t_{i+1}, 2 \leq i \leq n) \\ &= \prod_{i=2}^n \mathbb{P}(T_i - T_{i+1} \in \varepsilon dt_i - \varepsilon t_{i+1}) \\ &= \prod_{i=2}^n a_i \varepsilon e^{-a_i \varepsilon (t_i - t_{i+1})} dt_i. \end{aligned}$$

Integrating this last equation, it is not hard to show that as $\varepsilon \rightarrow 0$, $\mathbb{P}(T_2 < \varepsilon) \sim \varepsilon^{n-1} (\prod_{i=2}^n a_i) / (n-1)!$, a result that can also be obtained by taking the Laplace transform of T_2 and by applying a Tauberian theorem. This shows that as $\varepsilon \rightarrow 0$,

$$\begin{aligned} \mathbb{P}(T_i^\varepsilon \in dt_i, 2 \leq i \leq n) &= \mathbb{P}(\varepsilon^{-1}T_i \in dt_i, 2 \leq i \leq n | T_2 < \varepsilon) \\ &\sim \frac{\prod_{i=2}^n a_i \varepsilon e^{-a_i \varepsilon (t_i - t_{i+1})} dt_i}{\varepsilon^{n-1} \prod_{i=2}^n a_i / (n-1)!} \\ &\longrightarrow (n-1)! \prod_{i=2}^n dt_i, \end{aligned}$$

which terminates the proof. ■

Acknowledgements. The authors thank the *Center for Interdisciplinary Research in Biology* (CIRB, Collège de France) for funding.

References

- [BL03] Bertoin, J. and Le Gall, J.F. Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields*. 126, 261–288, 2003.
- [BL05] Bertoin, J. and Le Gall, J.F. Stochastic flows associated to coalescent processes. II. Stochastic differential equations. *Ann. Inst. H. Poincaré Probab. Statist.* 41, 307–333, 2005.
- [BL06] Bertoin, J. and Le Gall, J.F. Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.* 50, 147–181, 2006.

- [C75] Cannings, C. The latent roots of certain Markov chains arising in genetics: A new approach I. Haploid models. *Adv. in Appl. Probab.* 6, 260–290, 1975.
- [FV79] Fleming, W.H. and Viot, M. Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* 28 817–843, 1979.
- [K82] Kingman, J.F.C. The coalescent. *Stoch. Proc. Appl.* 13(3), 235–248, 1982.
- [K63] Knight, F. B. Random walks and a sojourn density of Brownian motion. *Trans. Amer. Math. Soc.* 109 5–86, 1963.
- [L08] Lambert, A. Population Dynamics and Random Genealogies. *Stoch. Models* 24 45–163, 2008.
- [L10] Lambert, A. The contour of splitting trees is a Lévy process. *Ann. Probab.* 38 348–395, 2010.
- [LG05] Le Gall, J.F. Random trees and applications. *Probability Surveys* 2, 245–311, 2005.
- [LP13] Lambert, A., Popovic, L. The coalescent point process of branching trees. *Ann. Appl. Prob.* 23(1) 99–144, 2013.
- [LUB16] Lambert, A., Uribe Bravo, G. The comb representation of compact ultrametric spaces. E-print *arXiv:1602.08246*, 2016.
- [P04] Popovic, L. Asymptotic genealogy of a critical branching process. *Ann. Appl. Prob.* 4, 2120–2148, 2004.
- [PW06] Pfaffelhuber P., Wakolbinger A. The process of most recent common ancestors in an evolving coalescent. *Stoch. Process. Appl.* 116, 183–1859, 2006.
- [PWW11] Pfaffelhuber P., Wakolbinger, A., Weisshaupt H. The tree length of an evolving coalescent. *Probab. Theory Relat. Fields.* 151, 529–557, 2011.
- [MS01] Möhle, M. and Sagitov, S. A classification of coalescent processes for haploid exchangeable population models. *Ann. Appl. Prob.* 29, 1547–1562, 2001.
- [RW87] Rogers, L.C.G. and Williams, D. Diffusions, Markov Processes, and Martingales. Wiley, New York, second edition, 1987.